

# Decimal Expansion Representation of Real Numbers

Fall 2016

Doğın Çömez

## Contents:

1. Introduction
2. Decimal Expansion
3. The Connection with the Ten-fold Map
4. The  $n$ -fold Map  $T_n$
5. The Full  $n$ -shift Space  $\Sigma_n$

## 1. Introduction

A real number can be represented in many ways. The most popular representation is by *decimal expansion*; i.e., expressing a real number as a geometric series with base 10. A modification of this representation is expressing a real number as a geometric series with base  $n$  for any positive integer  $n \geq 2$ , called as the  *$n$ -ary expansion*.

The decimal expansion of a real number provides a convenient method of representing rational and irrational numbers as well as approximations of irrational numbers by rational numbers. Furthermore, the process of obtaining the decimal expansion of a real number has close connections with dynamical systems and probability theory, which enables one investigate its properties in more depth. In this presentation I will briefly introduce the decimal expansion representation of real numbers, obtain some important properties, and then focus on the dynamical properties of the “ $n$ -ary” map and some of its consequences.

The concept of decimal expansion goes back to 8th century when Muhammad ibn-Musa Khwarizmi (c. 780-c. 850) introduced the arabic number system in his book who on algebra. In its numerous translations into latin during 12th century, one encounters occasional use of decimal notation. The regular use of the decimal point and decimal notation appears to have been introduced about 1585 by Flemish scientist Simon Stevinus (c. 1548-1620). It was Scottish mathematician John Napier (1550-1617), the inventor of logarithms in 1614, who first used and then popularized the decimal point to separate the whole part from the fractional part of a number.

When  $|r| < 1$ , the geometric series  $\sum_{k=0}^{\infty} r^k$  converges with the sum  $\frac{1}{1-r}$ . If  $r = \frac{1}{10}$  and  $a_i \in \{0, 1, \dots, 8, 9\}$ , for all  $i \geq 0$ , then the series

$$\sum_{k=0}^{\infty} \frac{a_k}{10^k}$$

converges to a real number by comparison test. The idea of decimal expansion is the process of obtaining the converse of this fact; namely, given a real number  $x \in [0, 1)$ , determining the sequence of integers  $0 \leq a_i \leq 9$  such that  $x = \sum_{k=0}^{\infty} \frac{a_k}{10^k}$ . There is nothing special about the

number 10; indeed, one can develop expansion of real numbers with respect to any other base  $n \in \mathbb{N}$  as well. In that case, the expansion is called as the  $n$ -ary expansion.

## 2. Decimal Expansion Representation

Let's look at the decimal expansion process closely. Given a real number  $x \geq 0$ .

(i) Let  $a_0 = [x]$ , where  $[\cdot]$  is the greatest integer function. So,  $x = a_0 + p_0$ , where  $0 \leq p_0 < 1$ . Therefore,  $a_0 \leq x < a_0 + 1$ , and hence  $10a_0 \leq 10x < 10a_0 + 9$ . Also,  $0 \leq 10p_0 < 10$ .

(ii) Now,  $10p_0 = a_1 + p_1$ , where  $a_1 = [10p_0]$  is an integer and  $0 \leq p_1 < 1$ . Thus,  $0 \leq a_1 \leq 9$ , and

$$a_1 \leq 10p_0 < a_1 + 1 \iff a_1 \leq 10(x - a_0) < a_1 + 1,$$

which implies that

$$a_0 + \frac{a_1}{10} \leq x < a_0 + \frac{a_1}{10} + \frac{1}{10}.$$

(iii) Similarly,  $10p_1 = a_2 + p_2$ , where  $a_2 = [10p_1]$  is an integer and  $0 \leq p_2 < 1$ . Thus,  $0 \leq a_2 \leq 9$ , and

$$\begin{aligned} a_2 \leq 10p_1 < a_2 + 1 &\iff a_2 \leq 10(10p_0 - a_1) < a_2 + 1 \\ &\iff a_2 \leq 10^2 p_0 - 10a_1 < a_2 + 1 \\ &\iff a_2 \leq 10^2(x - a_0) - 10a_1 < a_2 + 1, \end{aligned}$$

which implies that

$$a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} \leq x < a_0 + \frac{a_1}{10} + \frac{1}{10} + \frac{a_2}{10^2}.$$

Continuing in this process and letting  $10p_{k-1} = a_k + p_k$ , where  $a_k = [10p_{k-1}]$  and  $0 \leq p_k < 1$ , we obtain integers  $a_i \in \{0, 1, \dots, 9\}$ ,  $0 \leq i \leq k$ , such that

$$a_0 + \frac{a_1}{10} + \dots + \frac{a_k}{10^k} \leq x < a_0 + \frac{a_1}{10} + \dots + \frac{a_k}{10^k} + \frac{1}{10^k}.$$

Observe that if  $a_n = 0$  for some  $n$ , then  $a_i = 0$  for all  $i \geq n$ ; hence the process terminates. Otherwise the process continues indefinitely. In the case that the process terminates at the  $n$ th step, we have

$$(1) \quad x = a_0 + \frac{a_1}{10} + \dots + \frac{a_n}{10^n};$$

and when the process continues indefinitely, we have

$$(2) \quad x = a_0 + \frac{a_1}{10} + \dots + \frac{a_n}{10^n} + \dots = \sum_{k=0}^{\infty} \frac{a_k}{10^k}.$$

These expressions (1) and (2) are called the *decimal expansions* of the real number  $x$ , where (1) is finite and (2) is infinite, and are denoted by

$$x = a_0.a_1a_2 \dots a_{n-1}a_n, \quad \text{or} \quad x = a_0.a_1a_2 \dots a_n \dots, \quad \text{respectively.}$$

**Remarks.** 1. In the case that the process continues indefinitely, the decimal expansion process outlined above shows that, at any step  $n \geq 1$ , we have

$$a_0 + \frac{a_1}{10} + \dots + \frac{a_n}{10^n} \leq x < a_0 + \frac{a_1}{10} + \dots + \frac{a_n}{10^n} + \frac{1}{10}.$$

Hence, we obtain two sequences of finite decimals each converging to  $x$ , one increasing from below and the other decreasing from above. By repeating this procedure sufficiently many times, the decimal expansion of  $x$  can be obtained to any desired degree of accuracy.

2. In the case that the process continues indefinitely, there are two possibilities: either the sequence of terms  $\{a_i\}$  has no repetition, or it repeats, in the sense that, for some positive integers  $k$  and  $m$ ,  $a_k = a_{k+m}$ ; hence, the decimal expansion has the form

$$a_0.a_1a_2 \dots a_k a_{k+1} a_{k+2} \dots a_{k+m-1} a_k a_{k+1} \dots,$$

which we denote, for convenience, as

$$a_0.a_1a_2 \dots a_{k-1} \overline{a_k a_{k+1} a_{k+2} \dots a_{k+m-1}},$$

where  $\overline{a_k a_{k+1} a_{k+2} \dots a_{k+m-1}}$  denotes that the string of terms  $a_k, a_{k+1}, \dots, a_{k+m-1}$  repeats indefinitely.

3. If  $a_0.a_1a_2 \dots a_{k-1}a_k$  is a terminating decimal expansion, then it can also be rewritten as a repeating decimal expansion as

$$a_0.a_1a_2 \dots a_{k-1}a_k 000 \dots = a_0.a_1a_2 \dots a_k \overline{0}.$$

However, in order to distinguish such decimals from the decimals with nonzero repeating parts, we will call them as *finite decimal expansion*.

Observe that if  $x = a_0.a_1a_2 \dots a_k \overline{0} = \sum_{i=0}^k \frac{a_i}{10^i}$ ; hence  $x \in \mathbb{Q}$ . If  $x = a_0.a_1a_2 \dots a_{k-1} \overline{a_k a_{k+1} a_{k+2} \dots a_{k+m-1}}$ , then letting  $y = a_0.a_1a_2 \dots a_{k-1}$  and  $z = 0.a_k a_{k+1} a_{k+2} \dots a_{k+m-1}$ , we see that  $y, z \in \mathbb{Q}$ . Then,

$$\overline{a_k a_{k+1} a_{k+2} \dots a_{k+m-1}} = \frac{10^m}{10^m - 1} z;$$

and hence  $x = y + \frac{10^m}{10^m - 1} z \in \mathbb{Q}$ . Thus, any finite or repeating decimal expansion represents a rational number. The converse of this statement is also valid.

**Fact.1** If  $x \in \mathbb{Q}$ , then it has either finite or repeating decimal expansion.

**Proof.** For convenience, we will assume that  $x \in (0, 1)$ ; hence, we have  $a_0 = [x] = 0$ . Let  $x = \frac{r}{s}$ ,  $r, s \in \mathbb{Z}^+$ , and recalling the process of

obtaining decimal expansion of numbers above, we observe that  $p_0 = \frac{r}{s}$ , and

$$\begin{aligned} 10\left(\frac{r}{s}\right) &= a_1 + \frac{r_1}{s}, \text{ where } 0 \leq \frac{r_1}{s} < 1, \\ 10\left(\frac{r_1}{s}\right) &= a_2 + \frac{r_2}{s}, \text{ where } 0 \leq \frac{r_2}{s} < 1, \\ &\dots \\ 10\left(\frac{r_{k-1}}{s}\right) &= a_k + \frac{r_k}{s}, \text{ where } 0 \leq \frac{r_k}{s} < 1. \end{aligned}$$

Multiplying both sides by  $s$  at each step above, we see that this process is exactly the division algorithm:

$$\begin{aligned} 10r &= a_1s + r_1, \text{ where } 0 \leq r_1 < s, \\ 10r_1 &= a_2s + r_2, \text{ where } 0 \leq r_2 < s, \\ &\dots \\ 10r_{k-1} &= a_k + r_k, \text{ where } 0 \leq r_k < s. \end{aligned}$$

Notice that if one of  $r_k = 0$ , then  $r_i = 0$  for all  $i \geq k$ , i.e., the process stops and we have finite decimal expansion. For otherwise, since  $r_i \in \{1, 2, \dots, 9\}$ , there will be integers  $k < m$  such that  $r_m = r_k$ . Then, by division algorithm, we will have  $r_{m+i} = r_{k+i}$  for  $0 \leq i \leq m - k$ . In turn, we also have that  $a_{m+i} = a_{k+i}$  for  $0 \leq i \leq m - k$ ; hence the block  $a_k a_{k+1} \dots a_{m-1}$  repeats, which yields to repeating decimal expansion. ■

**Corollary.** A real number  $x$  is irrational if and only if it has a non-repeating decimal expansion.

**Remark.** Observe that, if  $x = a_0.a_1a_2 \dots a_n$ , i.e.,  $x$  has finite decimal expansion, then, by the division algorithm, we also have that

$$x = a_0.a_1a_2 \dots a_{n-1}(a_n - 1)999 \dots = a_0.a_1a_2 \dots a_{n-1}(a_n - 1)\bar{9}.$$

Thus, decimal expansion of a rational numbers need not be unique! This is an unpleasant situation.

**Question.** Can we circumvent this problem?

It turns out that the remedy is bringing some tools of dynamical systems into the process. Hence, dynamics comes to the rescue! Next we will discuss this in length and at the same time study an interesting type of dynamical systems.

### 3. The Connection with the Ten-fold Map

Let's look at the process of finding the (decimal) digits in decimal expansion process in an alternative way. Again, we will focus on  $x \in [0, 1)$ .

Partition  $[0, 1)$  into 10 equal subintervals  $[\frac{i}{10}, \frac{i+1}{10})$ ,  $0 \leq i \leq 9$ , and, if  $x \in [\frac{i}{10}, \frac{i+1}{10})$ , let  $a_1 = i$ . Then divide  $[\frac{i}{10}, \frac{i+1}{10})$  into 10 equal subintervals

$[\frac{i}{10} + \frac{j}{10^2}, \frac{i+1}{10} + \frac{j}{10^2})$ ,  $0 \leq j \leq 9$ , and, if  $x \in [\frac{i}{10} + \frac{j}{10^2}, \frac{i+1}{10} + \frac{j}{10^2})$ , let  $a_2 = j$ . Continuing in this manner, we obtain all the digits  $a_k$  such that

$$x = 0.a_1a_2a_3 \dots a_n \dots$$

At each step, this process is the same as multiplication of  $x$  by 10 and taking  $\text{mod } 1$ . Namely, repeated the action of the *ten-fold* map  $T$  on  $[0, 1)$ , where

$$T(x) = \begin{cases} 10x - 0 & \text{if } 0 \leq x < \frac{1}{10} \\ 10x - 1 & \text{if } \frac{1}{10} \leq x < \frac{2}{10} \\ \dots & \\ 10x - 9 & \text{if } \frac{9}{10} \leq x < 1. \end{cases}$$

Therefore,

$$\begin{aligned} a_1 = i_1 & \text{ if } \frac{i_1}{10} \leq x < \frac{i_1+1}{10} \iff Tx = 10x - i_1 \\ a_2 = i_2 & \text{ if } \frac{i_2}{10} \leq Tx < \frac{i_2+1}{10} \iff T^2x = 10(Tx) - i_2 \\ & \dots \\ a_k = i_k & \text{ if } \frac{i_k}{10} \leq x < \frac{i_k+1}{10} \iff T^kx = 10(T^{k-1}x) - i_k, \end{aligned}$$

and consequently,

$$\begin{aligned} x &= \frac{a_1}{10} + \frac{1}{10}Tx \\ &= \frac{a_1}{10} + \frac{1}{10}\left(\frac{a_2}{10} + \frac{1}{10}T^2x\right) \\ &\quad \dots \\ &= \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3} + \dots + \frac{a_k}{10^k}T^kx. \end{aligned}$$

**Remark.** It also follows from the arguments above that  $a_k = [10T^{k-1}x]$  when  $x = 0.a_1a_2a_3 \dots a_n \dots$ .

If  $x = \frac{m}{10^k}$  for some  $k, m \in \mathbb{Z}^+$ , then  $T^kx = 0$ ; and hence,  $x = \sum_{i=1}^{k-1} \frac{a_i}{10^i}$ . On the other hand, if  $T^kx \neq 0$ ,  $\forall k \in \mathbb{Z}^+$ , then taking limit as  $k \rightarrow \infty$ , we have  $x = \sum_{i=1}^{\infty} \frac{a_i}{10^i}$ . In particular,

- if  $x$  has finite (terminating) decimal expansion, then  $T^kx = 0$  eventually,
- if  $x$  has periodic infinite decimal expansion, then  $T^kx = x$  for some  $k$ , and
- if  $x$  has non-periodic infinite decimal expansion, then  $T^kx$  wanders in  $[0, 1)$ .

**Remark.** In this process no decimal expansion with infinite 999... occur. Hence, we settled the non-uniqueness problem!

If one looks at the behavior of  $T$  observed above more closely, much more interesting properties of  $T$  can be exhibited. For, let's define the *orbit of  $T$*  by  $O_T(x) = \{T^k x\}_{k=0}^{\infty}$ . Hence,

- a) if  $x \in \mathbb{Q} \cap [0, 1)$ , then  $O_T(x)$  is finite, and
- b) if  $x \in \mathbb{Q}^c \cap [0, 1)$ , then  $O_T(x)$  is finite,

which is a more descriptive characterization.

Next, observe that if  $x = \sum_{k=1}^{\infty} \frac{a_k}{10^k}$ , then  $Tx = \sum_{k=2}^{\infty} \frac{a_k}{10^k}$ . In other words

$$T(0.a_1a_2a_3 \dots a_k \dots) = 0.a_2a_3a_4 \dots a_{k+1} \dots,$$

so, the action of  $T$  on the decimal expansion of any real number is moving the decimal point one step right (right-shift)! These (and other) properties of  $T$  are special case of properties of a more general map, known as  $n$ -fold map and its properties as a dynamical system. That is why, in the rest of this note we will study this map in the dynamical system context.

#### 4. The $n$ -fold Map $T_n$

A topological dynamical system is a pair  $(X, T)$ , where  $X$  is a topological space and  $T : X \rightarrow X$  is a map compatible with the topology. Typically  $X$  is a metric space and  $T$  is a continuous map, or a map with finitely many discontinuities. Among such dynamical systems the  $n$ -fold (or  $n$ -ary) system has some special properties. In this section we will define and provide some basic features of the  $n$ -fold system that will pave way to an important class of symbolic dynamical systems called shift spaces.

Let  $X = [0, 1)$  with the usual Euclidean metric and, for any  $n \geq 2$ , let  $T_n : [0, 1) \rightarrow [0, 1)$  be defined by  $T_n(x) = nx(\text{mod } n)$ ,  $x \in [0, 1)$ . Hence,

$$T_n(x) = \begin{cases} nx & \text{if } 0 \leq x < 1/n \\ nx - 1 & \text{if } 1/n \leq x < 2/n \\ \dots & \\ nx - (n - 1) & \text{if } (n - 1)/n \leq x < 1. \end{cases}$$

The system  $([0, 1), T_n)$  is called the  $n$ -fold (or  $n$ -ary) system. When  $n = 2$  it is called the *doubling system (map)*. Some simple properties of the  $n$ -fold map are as follows:

1.  $T_n$  has  $n$  branches,  $i$ -th branch is maps the interval  $[\frac{i-1}{n}, \frac{i}{n})$  to  $[0, 1)$ ,  $1 \leq i \leq n$ .
2.  $T_n$  has periodic points of all orders. (A point  $x \in X$  is called a *periodic point of period  $n$*  of a system  $(X, T)$  if  $T^n x = x$  and  $n$  is the smallest such positive integer.) Periodic points of period

one are also called *fixed points*. 0 is the only periodic point for all  $n$ -fold maps. Indeed, an  $n$ -fold map has  $n - 1$  fixed points; for instance, binary map has only 0 as fixed point, 10-fold map has  $0, \frac{1}{9}, \frac{2}{9}, \frac{3}{9}, \dots, \frac{8}{9}$  as fixed points. Also, for the binary map, among others,  $1/3$  has period 2,  $4/7$  has period 3,  $1/5$  has period 4, . . . etc. Indeed, every rational point is a periodic point for the  $n$ -fold map. The converse is also valid.

3.  $T_n$  is *transitive*, that is, there exists  $x \in [0, 1)$  such that the set  $\{T^k x\}_{k=0}^{\infty}$  is dense in  $[0, 1)$ . In a dynamical system  $(X, T)$  the set  $O_T(x) = \{T^k x\}_{k=0}^{\infty}$  is called as the *orbit* of the point  $x \in X$ . Since rational points are those with periodic orbits by (3) above, irrational points are not periodic; hence, the transitive points of the  $n$ -fold map are among rationals. We will not exhibit an example of a transitive point of this system here, it'll be left to the next section when we investigate the shift spaces. If  $\overline{O_T(x)} = X$  for all  $x \in X$  in a dynamical system  $(X, T)$ , it is called a *minimal* system. Clearly,  $([0, 1), T_n)$  is not minimal.

For a dynamical system  $(X, T)$ , if  $X$  is endowed with a measure  $\mu$  (with the associated sigma algebra of measurable subsets of  $X$ , the triple  $(X, \mu, T)$  is also called a *measurable* dynamical system. If  $\mu(E) = \mu(T^{-1}E)$  for all measurable  $E$ , then we say that  $T$  *preserves*  $\mu$ , and  $\mu$  is  *$T$ -invariant*. If  $[0, 1)$  is endowed with the Lebesgue measure  $m$ , then  $T_n$  preserves  $m$ . For, it is enough to show the  $m(T_n(I)) = m(I)$  for any interval  $I \subset [0, 1)$ . If  $I = [a, b)$ , then  $T^{-1}I = \cup_{i=0}^{n-1} [\frac{a}{n} + \frac{i}{n}, \frac{b}{n} + \frac{i}{n})$ . Hence,

$$m(T_n^{-1}[a, b)) = \sum_{i=0}^{n-1} m([\frac{a}{n} + \frac{i}{n}, \frac{b}{n} + \frac{i}{n})) = b - a = m([a, b)).$$

$m$  is not the only  $T_n$ -invariant measure on  $[0, 1)$ , indeed there are uncountably many measures  $\mu$  on  $[0, 1)$  (equivalent to  $m$ ) that are  $T_n$ -invariant. Again, we will exhibit such measure in the next section. If  $\mu$  is the only  $T$ -invariant measure of a dynamical system  $(X, \mu, T)$ , then the system is called *uniquely ergodic*. Thus,  $([0, 1), m, T_n)$  is not uniquely ergodic.

A map  $T$  of a dynamical system  $(X, \mu, T)$  is called *ergodic* if any measurable set  $E$  with  $E = T^{-1}E$  (called  *$T$ -invariant set*) has either measure 0 or full measure. It turns out that  $T_n$  is ergodic. In order to prove the ergodicity of  $T_n$ , we need a few technicalities.

A class  $\mathcal{C}$  of subintervals of  $[0, 1)$  is called a *covering class* for  $[0, 1)$  if every subinterval of  $[0, 1)$  is a countable disjoint union of elements from  $\mathcal{C}$ .

**Lemma.** (Knopp's Lemma) Let  $B \subset [0, 1)$  be a measurable set and  $\mathcal{C}$  be a covering class for  $[0, 1)$ . If

$$(*) \quad \forall A \in \mathcal{C}, \quad m(A \cap B) \geq \gamma m(A),$$

where  $\gamma > 0$  is independent of  $A$ , then  $m(B) = 1$ .

**Proof.** We will prove the statement by contradiction. Assume that  $m(B) < 1$ , i.e.,  $m(B^c) > 0$ . Since  $B$  is a measurable set,  $B = C \cup D$ , where  $C$  is a Borel set with  $m(B) = m(C)$  and  $m(D) = 0$ . Thus  $m(C^c) > 0$  as well. Now, given  $\epsilon > 0$ , there a set  $E_\epsilon$  which is a finite disjoint union of open intervals in  $[0, 1)$  such that  $m(C^c \Delta E_\epsilon) < \epsilon$ . Hence,  $E_\epsilon$  is a countable disjoint union of elements from  $\mathcal{C}$ . Therefore, by (\*), it follows that  $m(C \cap E_\epsilon) \geq \gamma m(E_\epsilon)$ . Then,

$$m(C^c \Delta E_\epsilon) \geq m(C \cap E_\epsilon) \geq \gamma m(E_\epsilon) \geq \gamma m(C^c \cap E_\epsilon) > \gamma(m(E_\epsilon) - \epsilon).$$

This implies that  $\gamma(m(E_\epsilon) - \epsilon) < m(C^c \Delta E_\epsilon) < \epsilon$ . Therefore, we must have  $\gamma m(C^c) < \epsilon + \gamma\epsilon$ . Since  $\epsilon$  is arbitrary, this implies that  $m(C^c) = 0$ , contradiction. ■

**Fact.1**  $T_n$  is ergodic.

**Proof.** We will prove the that  $T_2$  is ergodic for simplicity, ergodicity of  $T_n$  follows the same lines. Let  $B$  be a  $T_2$ -invariant measurable subset of  $[0, 1)$ . Consider dyadic intervals  $D_{n,k} = [\frac{k}{2^n}, \frac{k+1}{2^n})$ , where  $n$  is a positive integer and  $k = 0, 1, 2, \dots, 2^n - 1$ . Then the collection  $\mathcal{C}$  of all dyadic intervals is a covering class for  $[0, 1)$ . Also,  $m(D_{n,k}) = 2^{-n}$  and  $T_2^n(D_{n,k}) = [0, 1)$  for each  $k = 0, 1, 2, \dots, 2^n - 1$ . Also, it follows by induction that, for any measurable set  $A$ ,

$$m(T_2^{-n} A \cap D_{n,k}) = 2^{-n} m(A) = m(A) m(D_{n,k}), \quad k = 0, 1, 2, \dots, 2^n - 1.$$

Since  $B$  is  $T_2$ -invariant, his implies that

$$m(B \cap D_{n,k}) = m(B) m(D_{n,k}), \quad \text{for all } n > 0, \quad k = 0, 1, 2, \dots, 2^n - 1.$$

If  $\gamma = m(B) > 0$ , then it follows that  $m(B \cap C) \geq \gamma m(C)$  for any  $C \in \mathcal{C}$ . Hence, by Knopp's Lemma,  $m(B) = 1$ , i.e.,  $T_2$  is ergodic. ■

**Remark.** A map  $T$  of a dynamical system  $(X, \mu, T)$  is called *totally ergodic* if  $T^m$  is ergodic for all  $m \geq 1$ . Since  $T_n$  is ergodic and since  $T_n^m = T_{nm}$ , it follows that  $T_n$  is totally ergodic.

A map  $T$  of a dynamical system  $(X, \mu, T)$  is called *mixing* if for all measurable sets  $A$  and  $B$ ,  $\lim_{n \rightarrow \infty} \mu(T^{-n} A \cap B) = \mu(A)\mu(B)$ . Observe that in the proof of ergodicity of  $T_2$  we have shown that, for any measurable set  $A$ ,

$$m(T_2^{-n} A \cap D_{n,k}) = m(A) m(D_{n,k}).$$

Since dyadic intervals area generating  $\sigma$ -algebra for measurable sets, it follows that

$$\lim_n m(T_2^{-n} A \cap B) = m(A) m(B),$$

for all measurable sets  $A$  and  $B$ . Thus  $T_2$  is mixing. By the same proof (adapted to  $n$ -adic intervals), it follows that  $T_n$  is mixing.

There is an intermediate property between ergodicity and mixing. A map  $T$  of a dynamical system  $(X, \mu, T)$  is called *weakly mixing* if for all measurable sets  $A$  and  $B$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(T^{-k}A \cap B) - \mu(A)\mu(B)| = 0.$$

By the Fact .2 below we have that

$$\text{mixing} \Rightarrow \text{weak mixing} \Rightarrow \text{ergodicity},$$

hence,  $T_n$  is weakly mixing.

**Fact.2** For a dynamical system, mixing  $\Rightarrow$  weak mixing  $\Rightarrow$  ergodicity.

**Proof.** By the Ergodic Theorem (see [5])  $T$  is ergodic if and only if for all measurable  $A, B$ ,

$$(1) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}A \cap B) = \mu(A)\mu(B).$$

Also, it is known that, for any sequence  $(a_n)$  of real numbers,

$$(2) \quad \lim a_n = 0 \Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n |a_k| = 0 \Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n a_k = 0.$$

From (1) and (2) the assertion follows. ■

Two (topological) dynamical systems  $(X, T)$  and  $(Y, S)$  are called *conjugate*, and denoted by  $X \equiv Y$ , if there exists a homeomorphism  $\phi : X \rightarrow Y$  such that  $\phi \circ T = S \circ \phi$ . Let  $\Sigma_2 = \{0, 1\}^{\mathbb{N}}$ , the space of all sequences of 0's and 1's. Define a metric  $d : \Sigma_2 \times \Sigma_2 \rightarrow \mathbb{R}^+$  by

$$d((a_k), (b_k)) = \sum_{k=1}^{\infty} \frac{|a_k - b_k|}{2^k}, \quad (a_k), (b_k) \in \Sigma_2.$$

Then  $(\Sigma_2, d)$  is a metric space. Since  $\{0, 1\}$  is a compact metric space, by Tychonoff's Theorem, so is  $(\Sigma_2, d)$ . Now, define a map, called as *the (left) shift map*,  $\sigma : \Sigma_2 \rightarrow \Sigma_2$  by  $\sigma((a_k)) = (a_{k+1})$ . The dynamical system  $(\Sigma_2, \sigma)$  is called the *2-shift space* (or *shift space*).

**Remarks.** 1. The metric space  $(\Sigma_2, d)$  has topological dimension zero.

2.  $\Sigma_2$  is homeomorphic to Cantor set.

3. The shift map  $\sigma : \Sigma_2 \rightarrow \Sigma_2$  is two-to-one map (in the case  $\sigma : \Sigma_n \rightarrow \Sigma_n$ , it is an  $n$ -to-one map).

4. If  $\Sigma_n^* = \{0, 1, \dots, n-1\}^{\mathbb{Z}}$ , then the map  $\sigma : \Sigma_n^* \rightarrow \Sigma_n^*$  is a one-to-one, onto and homeomorphism.

It is well-known that any  $x \in [0, 1)$  can be represented in *binary expansion*

$$x = \sum_{k=1}^{\infty} \frac{a_k}{2^k} = 0.a_1a_2a_3 \dots a_k \dots,$$

where  $a_k \in \{0, 1\}$ . This representation is unique except for countably many (rational) numbers, i.e., for those with binary expansion of the form  $*1000 \dots 0 \dots$ , which also has binary expansion of the form  $*0111 \dots 1 \dots$ . One can remedy this non-uniqueness by accepting the second expansion only. Observe that, for any  $x \in [0, 1)$  with binary expansion  $(a_k)_k$ ,

$$\begin{aligned} T_2(x) &= 2\left(\sum_{k=1}^{\infty} \frac{a_k}{2^k}\right) \pmod{2} = \sum_{k=2}^{\infty} \frac{a_k}{2^{k-1}} \\ &= \sum_{k=1}^{\infty} \frac{a_{k+1}}{2^k} = 0.a_2a_3a_4 \dots a_{k+1} \dots \end{aligned}$$

Thus,  $T_2(0.a_1a_2a_3 \dots a_k \dots) = 0.a_2a_3a_4 \dots a_{k+1} \dots$ . Now, if we define the map  $\phi : [0, 1) \rightarrow \Sigma_2$  by  $\phi(x) = (a_k)_{k=1}^{\infty}$ , where  $x = 0.a_1a_2a_3 \dots a_k \dots$ , then  $\phi$  is a bijection and a continuous map with continuous inverse. Furthermore,

$$\begin{aligned} \phi(T_2x) &= \phi(0.a_2a_3a_4 \dots a_{k+1} \dots) = (a_2a_3a_4 \dots a_{k+1} \dots) \\ &= \sigma(a_1a_2a_3 \dots a_k \dots) = \sigma(\phi(0.a_1a_2a_3 \dots a_k \dots)) \\ &= \sigma(\phi(x)). \end{aligned}$$

Hence,  $\phi \circ T_2 = \sigma \circ \phi$ , i.e.,  $([0, 1), T_2)$  and  $(\Sigma_2, \sigma)$  are conjugate spaces. Notice that, for any  $m \geq 1$ ,  $T_2^m(x) \equiv \sigma^m(a_k)$ , where  $x = 0.a_1a_2a_3 \dots a_k \dots$ .

Let  $\Sigma_n = \{0, 1, 2, \dots, n-1\}^{\mathbb{N}}$ , the space of all sequences of letters from the alphabet  $\mathcal{A} = \{0, 1, 2, \dots, n-1\}$ . As in the case of  $\Sigma_2$ , this space is also a compact metric space with the metric

$$d((a_k), (b_k)) = \sum_{k=1}^{\infty} \frac{|a_k - b_k|}{n^k}, \quad (a_k), (b_k) \in \Sigma_n.$$

If we define the shift map as above, the system  $(\Sigma_n, \sigma)$  is called the *n-shift space* (or *shift space*). Similarly, it follows that  $([0, 1), T_n)$  and  $(\Sigma_n, \sigma)$  are conjugate spaces with the same conjugacy map  $\phi$ .

## 5. The Full $n$ -shift Space $\Sigma_n$

In this section we will investigate the (symbolic) dynamical system  $(\Sigma_n, \sigma)$ . Since it is conjugate to  $([0, 1), T_n)$ , we will also obtain some (topological) properties of the  $n$ -fold map. In what follows, for simplicity, we will only deal with  $\Sigma = \Sigma_2$ .

As in the case of  $n$ -fold map,  $\Sigma$  also has periodic point of all orders. Indeed, it is very easy to exhibit such points: any point  $\mathbf{x} \in \Sigma$  (i.e., any periodic sequence) of the form

$$\mathbf{x} = (a_1, a_2, \dots, a_m, a_1, a_2, \dots, a_m, a_1, \dots)$$

is a periodic point of period  $m$ . Since  $a_i \in \{0, 1\}$ , there are  $2^{m-1}$  periodic points of period  $m$  (we omit the point  $\mathbf{1} = (1, 1, 1, \dots)$  since it is not in  $\Sigma$ ). Consequently, there are only countable many periodic points of  $\Sigma$  and they form a dense set.

Since the  $n$ -fold map  $T_2$  is transitive, so is  $\sigma$ . On the other hand it is much easier to exhibit a point with dense orbit for  $\Sigma$  than for  $T_2$ : let

$$\mathbf{x} = (01000110110000010100111001011101110000\dots).$$

Observe that  $\mathbf{x}$  contains strings of all possible combinations of 0's and 1's in any length.

*Claim.*  $\overline{O_\sigma(\bar{x})} = \Sigma$ . For, let  $\mathbf{y} = (a_1, a_2, \dots, a_n, \dots) \in \Sigma$  be arbitrary and let  $\epsilon > 0$  be given. Pick  $n > 1$  such that  $\frac{1}{2^n} < \epsilon$ . Since  $\mathbf{x}$  contains all possible strings of 0's and 1's of length  $n$ , applying  $\sigma$  to  $\mathbf{x}$  sufficiently many times, say  $N$ , we have  $\sigma^N \mathbf{x} = (a_1, a_2, \dots, a_n, \dots)$ . Then  $\sigma^N \mathbf{x}$  and  $\mathbf{y}$  agree at the first  $n$  coordinates; hence,

$$d(\sigma^N \mathbf{x}, \mathbf{y}) \leq \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^n} < \epsilon.$$

Thus, the orbit of  $\mathbf{x}$  is dense in  $\Sigma$ ; hence,  $\Sigma$  is transitive. Notice that since it has many periodic points,  $\Sigma$  is not minimal.

**Remark.** The set of transitive points of  $\Sigma$  form a dense  $G_\delta$ -set in  $\Sigma$ .

These observations suggest that instead of studying  $T_n$ , one might study  $\Sigma_n$  and transfer the properties that are invariant under conjugacy (i.e., topological invariants) to the former. How about measurable properties? For, one needs to put a (suitable) measurable structure on  $\Sigma$  and show that  $([0, 1], T_2)$  and  $(\Sigma, \sigma)$  are measure theoretically conjugate (i.e., *isomorphic*) as well.

Let  $[a]_k := \{(x_n) \in \Sigma : x_k = a\}$ , where  $a \in \{0, 1\}$ . In general, let

$$[a_0, a_1, \dots, a_m]_k := \{(x_n) \in \Sigma : x_k = a_0, x_{k+1} = a_1, \dots, x_{k+m} = a_m\},$$

where  $a_i \in \{0, 1\}$ ,  $0 \leq i \leq m$ . Notice that

$$[a_0, \dots, a_m]_k = \{0, 1\} \times \dots \times \{0, 1\} \times \{a_0\} \times \{a_1\} \times \dots \times \{a_m\} \times \{0, 1\} \times \dots;$$

hence, such subsets of  $\Sigma$  are called the (basic) *cylinder sets*.

**Fact.3** The cylinder sets are both open and closed, and the collection  $\mathcal{C}$  of cylinder sets forms a countable base for the topology of  $\Sigma$ .

**Proof.** Let  $C = [a_0, \dots, a_m]_k$ . First, we will show that  $C$  is open. For, if  $\mathbf{x} \in C$ , then  $\mathbf{x} = (x_1, x_2, \dots, x_{k-1}, a_0, a_1, \dots, a_m, x_{k+m+1}, \dots)$ . Let  $\epsilon = \frac{1}{2^{k+m+1}}$ . Then

$$\begin{aligned} \mathbf{y} \in B(\mathbf{x}, \epsilon) &\iff \sum_{i=1}^{\infty} \frac{|x_i - y_i|}{2^i} < \frac{1}{2^{k+m+1}} \\ &\iff x_i = y_i \text{ for } 1 \leq i \leq n + k + 1, \end{aligned}$$

which implies that  $y_i = a_i$  for  $k \leq i \leq n + k + 1$ ; equivalently,  $\mathbf{y} \in C$ . Therefore,  $B(\mathbf{x}, \epsilon) \subset C$ ; i.e.,  $C$  is open.

Nest, let  $(\mathbf{x}^n) \subset C$  be a sequence such that  $\mathbf{x}^n \rightarrow \mathbf{x}$ . So, for all  $\epsilon > 0$ ,  $\exists N > 0$  such that  $r \geq N$  implies  $d(\mathbf{x}^r, \mathbf{x}) < \epsilon$ . Thus, for  $r \geq N$ , we must have  $\sum_{i=1}^{\infty} \frac{|x_i - x_i^r|}{2^i} < \epsilon$ . Therefore, if  $\epsilon$  is small enough, this implies that for all except finitely many  $r$ 's we have  $x_i^r = x_i$ ,  $1 \leq i \leq n + k + 1$ . Since such  $\mathbf{x}^r$ 's are in  $C$ , for  $r \geq N$ , we have

$$\mathbf{x} = (x_i) = (*, *, \dots, *, a_0, a_1, \dots, a_m, *, * \dots) \in C.$$

Therefore,  $C$  is closed.

Clearly, there are countably many cylinder sets of the form  $C$ . Given any  $\mathbf{x} \in \Sigma$  and  $\epsilon > 0$ , there is a cylinder set  $C$  such that  $\mathbf{x} \in C \subset B(\mathbf{x}, \epsilon)$ . Hence the collection  $\mathcal{C}$  forms a countable base for the topology of  $\Sigma$ . ■

**Remarks.** 1. By the Fact above, every open set is a countable union of cylinder sets. Since the cylinder sets are clopen and generate the topology of  $\Sigma$ , the metric space  $\Sigma$  has topological dimension zero.

2. The collection  $\mathcal{C}$  also forms a covering class for  $\Sigma$ ; namely, (i)  $\emptyset \in \mathcal{C}$ , and (ii)  $\forall A_i \subset \Sigma$ ,  $\exists \{C_n\} \in \mathcal{C}$  such that  $A \subset \cup_n C_n$ .

Since  $\mathcal{C}$  is a covering class for  $\Sigma$ , any non-negative  $\mathbb{R}^\#$ -valued set function  $\lambda$  with domain  $\mathcal{C}$  satisfying  $\lambda(\emptyset) = 0$  gives rise to an outer measure  $\mu^*$  by

$$\mu^*(A) = \inf \left\{ \sum_n \lambda(C_n) : \{C_n\} \subset \mathcal{C}, A \subset \cup_n C_n \right\}.$$

This, in turn, defines a measure on the Borel  $\sigma$ -algebra of subsets of  $\Sigma$ . So, let's construct such a measure.

Let  $\mathbf{p} = (p_0, p_1)$  such that  $0 \leq p_0, p_1 \leq 1$  and  $p_0 + p_1 = 1$ . Define  $\nu(\{0\}) = p_0$  and  $\nu(\{1\}) = p_1$ . Then  $\nu : \{0, 1\} \rightarrow \mathbb{R}^+$  is a probability measure (on  $\{0, 1\}$ ). Now, on the collection of cylinder sets  $\mathcal{C}$  define a set function  $\lambda : \mathcal{C} \rightarrow \mathbb{R}^+$  by

$$\lambda([a_0, a_1, \dots, a_m]_k) = \prod_{i=0}^m p_{a_i}, \quad a_i \in \{0, 1\}.$$

Let  $\mu$  be the restriction of the outer measure  $\mu^*$  generated by  $\lambda$  to the Borel  $\sigma$ -algebra of subsets of  $\Sigma$ . Then, by Kolmogorov Extension Theorem, the measure  $\mu$  is uniquely defined. Notice that one can construct

uncountable many such measures, called *Bernoulli measures*, and denoted by  $\mu_{\mathbf{p}}$ . Thus,  $(\Sigma, \mathcal{B}, \mu_{\mathbf{p}})$  is a probability measure for each choice of  $\mathbf{p}$ .

Now, observe that if  $C = [a_0, a_1, \dots, a_m]_k$  is a cylinder set, then

$$\begin{aligned}\sigma^{-1}(C) &= \{(x_n) \in \Sigma : \sigma(x_n) \in C\} \\ &= \{(x_n) \in \Sigma : x_{k+i+1} = a_i, 0 \leq i \leq m\},\end{aligned}$$

which is also a cylinder set. Thus,  $\sigma$  is a measurable function on cylinder sets; hence, it is measurable on  $\mathcal{B}$ . Furthermore,

$$\mu_{\mathbf{p}}(C) = \prod_{i=0}^m p_{a_i} = \mu_{\mathbf{p}}(\sigma^{-1}(C));$$

hence,  $\sigma$  preserves  $\mu_{\mathbf{p}}$ . It follows that the system  $(\Sigma, \mathcal{B}, \mu_{\mathbf{p}}, \sigma)$  is a measure preserving dynamical system!

### References

- [1 ] K. Dajani and C. Kraaikamp, **Ergodic Theory of Numbers**, Carus Math. Monographs #29, MAA Publications, 2002. ISBN 0-88385-034-6
- [2 ] I. Niven, **Irrational Numbers**, Carus Math. Monographs #11, MAA Publications, 1956.

**Notes.** [2] is one of the best sources for general information on rational and irrational numbers, decimal expansion and continued fraction expansion. Furthermore, it contains very interesting features of the irrational numbers.

[1] is the best and most comprehensive source on symbolic dynamics and the ergodic theory of number representation processes. It contains proofs of some major results on  $\beta$ -shifts and Gauss map while providing many relatively recent results on number theoretical, measure theoretical and topological properties of these transformations.